

DOCUMENT RESUME

ED 238 926

TM 840 025

AUTHOR Long, Hamish A.
 TITLE The Scottish Development-Towards a System of Positive Grade Related Assessment and Reporting for All.
 PUE DATE 5 Mar 83
 NOTE 17p.; Paper presented at the Annual Meeting of the International Association for Educational Assessment (9th, Blantyre, Malawi, June 13-20, 1983).
 PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS *Academic Achievement; Curriculum Development; *Educational Assessment; Elementary Education; English; Foreign Countries; Learning Problems; Mathematics; *Measurement Techniques; *Profiles; Sciences; *Slow Learners; Student Evaluation
 IDENTIFIERS *Scotland

ABSTRACT

Implementation of the Munn/Dunning proposals in Scotland for assessment and certification for all depend on the results of a three year program of which a major part has been the development of new courses and appropriate assessment instruments in Foundation English, Mathematics, and Science for the lowest 30 percent of the ability range. The main thrust of the development in assessment has been to define grade-related criteria applicable to both external and internal assessment. In this way the external examining agency can monitor the schools' internal assessment. Comparisons will be carried out in 1983 on a mode by mode basis. This will enable an accurate profile of pupil achievement since criteria are stated in positive subject attainment terms. The work is now being extended to the General and Credit levels in eleven subjects with the definition, over the next year, or courses, assessment instruments and grade-related criteria for seven articulated award grades spanning the whole ability range. (Author/PN)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ABSTRACT

THE SCOTTISH DEVELOPMENT - TOWARDS A SYSTEM OF POSITIVE GRADE RELATED ASSESSMENT AND REPORTING FOR ALL

Dr H A Long
Scottish Examination Board
~~8-12-82~~ 3/5/83

Implementation of the Munn/Dunning proposals in Scotland for assessment and certification for all depend on the results of a three year programme of which a major part has been the development of new courses and appropriate assessment instruments in Foundation English, Mathematics and Science for the lowest 30% of the ability range in S3 and S4.

It has been shown that pupils in this range can cope with external examinations contrary to prior expectation. The main thrust of the development in assessment, however, has been to define grade related criteria applicable to both external and internal assessment. Subjects have been subdivided into natural modes or skill areas: in English, for example, into Writing, Reading, Listening and Talking. Teachers have assisted in the development of criteria written in subject terms to assess these modes internally. Wherever possible a comparable external assessment instrument has been designed to assess each mode against the same criteria. In this way the external examining agency can monitor the schools' internal assessment. Comparisons will be carried out in 1983 on a mode by mode basis. This will enable an accurate profile of pupil achievement since criteria are stated in positive subject attainment terms. Rank orders of pupils are not prepared since marking is by direct grading against the criteria. Where certain tests naturally produce an interval scale, such as objective tests, pre-determined cut-off points are derived using the methods of Nedelsky or Angoff, or by reference to a calibrated item bank. Moderation of internal assessment by the traditional method of applying standardising factors is not applicable. The work is now being extended to the General and Credit levels in eleven subjects with the definition, over the next year, of courses, assessment instruments and grade related criteria for seven articulated award grades spanning the whole ability range.

Many problems remain but the development is aimed at their resolution.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. A. Long

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A PAPER FOR THE 9TH INTERNATIONAL CONFERENCE OF IAEA,
BLANTYRE, MALAWI, JUNE 1983

THE SCOTTISH DEVELOPMENT
TOWARDS A SYSTEM OF POSITIVE GRADE RELATED ASSESSMENT
AND REPORTING FOR ALL

Dr H A Long
S E B Dalkeith Scotland 3 May 1983

Dr Dockrell has presented an admirable outline of the background to the development currently being pursued - a development undergoing feasibility study when the original abstract of this paper was prepared some months ago but now being consolidated for full operational implementation with certification in 1986 by Government directive.

In Dr Dockrell's paper presented in 1977 to the Third Annual Conference of the IAEA it was indicated that reforms to the Scottish Examination system would be shortly forthcoming but that these reforms were in the future. Well at last the future has become the present in Scotland though there are some critics who would say the future is now the past, that the reforms are too late, that they were relevant to the last decade but not to the coming one and that they are not radical enough. Well I shall leave it to you to judge from my description of the evolution of the system which follows.

I shall refer not only to the subject of English but also to the subjects of Mathematics and Science, where appropriate, since these subjects were considered to be the primary essentials for all pupils regardless of ability and were developed in parallel as the first phase of the development. The emphasis throughout the development as you will have realised was not primarily on assessment nor on certification but on the provision of appropriate courses for the less able to make amends for the second rate provision offered to them to date.

A further guiding principle was that, while new courses would be developed, the courses themselves should articulate in such a way that pupils would not be compartmentalised, never to escape from a course once started and labelled for all time as being of certain ability in a subject prior to receiving education within that discipline. It was essential therefore that courses had to articulate between levels to allow pupils the freedom of vertical movement between courses until their most appropriate level was found. Finally assessment had to be given proper prominence both for diagnostic and formative purposes as well as summative for national certification. Certification indeed was seen as the key issue, the very cornerstone, of this development since past experience of schemes in previous decades to provide courses for the least able foundered one after the other through

lack of national recognition and certification. They were sometimes praised, often applauded, but when it came to public and teacher recognition, even pupil recognition, they were inevitably classed as third rate and not worth the trouble. Attempts to alleviate this through the provision of local or regional certificates had little impact.

The new development on the other hand has shaped courses and assessment into a completely new form for Scotland from those with which we in that country are familiar, accustomed and occasionally comfortable. In the first place, as you have heard, the courses had to be such that all pupils had to have a real chance of success. This would have to be mirrored by the assessment which would have to move as far as possible towards criterion referencing (further research on this continues) and national reporting had to be in positive subject terms.

The Dunning scheme suggested three articulated levels of courses and attainment (Credit, General and Foundation) with internal syllabus components and internal/external assessment in a 25%/75% weighting, the internal assessment to be scaled statistically on the external assessment. The assessments would then be aggregated and reported as a single award in terms of Credit, General 1 to 3, or Foundation. Further, provision would have to be made within the assessment to allow for enhancement to a grade above that on which the pupil was presented, or for compensation in which an award could be given on a lower grade than that of presentation. When the Government decided to proceed with the development of a possible system in 1980 they made a significant modification to the scheme in that internal syllabus development and internal assessment (except in the cases of practical or oral work) were to be confined to Foundation level. Credit and General grades were to be assessed by external assessment only. This, as you can imagine, caused a number of problems in attempting to provide articulated courses and assessment with enhancement and compensation when two of the adjoining levels were constructed to different frameworks, one with internal syllabus and assessment, one without. Since that time a number of concessions have been made which have eased the problem to some extent, and the Government has indicated that it will further review the situation in the light of experience. A further modification resulted from preliminary work done by the Scottish Education Department Inspectors who wrote provisional guidelines for English, Mathematics and Science and started the pilot work in schools at Foundation level. Prior to the Examination Board involvement they had decided that pupils should receive direct gradings and that rank orders and extended scales would not be used. In this they had the support of the pilot school teachers who were anxious to take up the challenge of grading competence rather than having to conform to what has been described as the Board's 'oppressive, iniquitous and fossilised norm-referenced system'. One immediate consequence

which can easily be seen is that the Dunning proposal for statistical scaling using external exam scores was no longer feasible.

This then was the revised starting grid for the development work we have been involved in over the last three years. My aim is to show you briefly how, through assessment trials and a consideration of other factors, the scheme was modified, some would argue improved, in an effort to make its operation possible as well as testing its feasibility, and to illustrate the many obstacles, hurdles, pitfalls and traps that had to be overcome, jumped or avoided by adopting modified strategies or philosophies through the course of the development phase. Feeding into this, I must make mention of the support, views and opinions of the pilot school teachers, all of which were considered, many of which were taken into account in devising the new system.

Now let us consider the assessment trials. As has been indicated the pilot work was carried out at Foundation level where it was felt the need was greatest, experience was least, and the problems most numerous. Bearing in mind the following constraints:

- (i) that all pupils should have a good chance of success;
- (ii) that provision should be made for articulation with a higher level of syllabus and award;
- (iii) that there should be a move towards criterion referencing;

it was decided that pupils would be graded in only three ways. These were Enhanced; Foundation; Course Completed. The majority of children following the pilot Foundation course in a subject should receive a Foundation award; those who clearly demonstrated a higher standard of achievement than the course demanded should receive an Enhanced award (i.e. an award comparable to that proposed for achievement on the General); those who failed to reach a minimum competence level for Foundation but had however attempted the course with honest endeavour would receive an award of Course Completed. The few (hopefully) who through truancy failed to attend classes or failed to make any reasonable effort would be placed in a 4th category: No Award. It was therefore essential to ensure that the assessment both internal and external attempted to meet the requirements outlined above. The courses too were evolving throughout this period. In English, for example, a conscious decision was made that pupils should be assessed in Reading, Writing, Listening and Talking. Mathematics was split into 6 assessable elements including Basic Skills, Data Processing, Order, Communication, Problem Solving and Practical Skills. Science took a different route by identifying assessment instruments rather than elements. These included for example Practical Investigations, Essays, Open-book tests, End of topic tests and so on. In Science I should also mention that a research project was commenced at this time to identify specific practical

skills to be tested and short items to permit teachers to carry out the assessment of these skills. At present some 52 objectives within six skill areas have been identified and several hundred items produced and pretested for teacher use.

Returning to the first year of the development, and considering internal assessment first, teachers were asked to submit grades on the 3-point scale outlined for each candidate for each domain or assessable element, or in the case of science for each instrument. For the external assessment much endeavour was devoted to devising external papers which covered the aspects graded by the internal assessment, which would yield high mean scores and which would be publicly acceptable. Needless to say many problems were encountered. Let me list just a few of them:

- (i) The initial criteria for assessment of the domains were undergoing constant rewriting; teachers found the task of grading their pupils very difficult particularly in aspects such as talking in English or practical skills in Maths where there was little prior experience of such assessment for summative purposes.
- (ii) Statistical scaling, as I have already mentioned was ruled out by the use of three point scales and the absence of rank orders.
- (iii) All internal assessments were regraded by a team of examiners, comprised mainly of a sample of the pilot school teachers themselves. The regrading exercise proved extremely difficult. In the case of Mathematics it was abandoned because the examiners could not disentangle the different assessable elements within the school course work which was submitted for regrading. Other problems arose through the variation and unsuitability of evidence submitted by schools. The interpretation of the 'first-attempt' criteria was difficult. In science the examiners modified the criteria for their regrading which prevented really useful comparison with the teachers' gradings on the original criteria. In English, teachers submitted tapes of talk in every conceivable form which lasted, for some pupils, for 30 minutes and, for others, a few seconds of speech within a class discussion or ten minutes of teacher questioning.

All was apparent confusion. In Science, again, the endless array of assessment instruments often unsupported by original texts, tests or marking schemes made the task nigh impossible. In English, again, evidence for listening and reading was second hand in that it consisted of written product from reading or listening.

- (iv) At the same time the implications of a 3 point scale finally came home when it was realised that without orders of merit there was no way, no matter how successful the regrading, that this formed a practicable method

of moderation in an operational situation. Without an order of merit no standardising factors could be applied, nor could interface candidates be identified. Regrading would have to become double marking and there was no way in which the economic or resource implications of such a step could be taken on board. Besides if everything were to be regraded what was the point of the internal assessment in the first instance? And teachers wanted nothing to do with orders of merit since this was seen as a mere reversion to norm-referencing.

All, however, was not gloom and despondency. For a start the external assessments were completed successfully by many candidates with high mean scores in spite of the prophecy of many that few of these least able pupils would be able to read the papers let alone answer them. (Side issues like this, incidentally were investigated with some pupils having the papers read to them; differences were not significant as far as performance was concerned). On the whole the impression was favourable. The major difficulty here was in allocating grades and fixing cut-off points for the external tests using such methods as Nedelsky and Angoff. This has been pursued over successive years and while there is no common solution the matter seems nearer resolution using a system based on grade related criteria rather than on arbitrarily chosen formulae. And too, to return to the internal work of the schools, there was much that was heartening. There was no doubt that the pupil work was far superior in many instances and taken as a whole to work produced by this range of pupils in previous circumstances. The courses seemed to be working, the pupils and teachers were more motivated. If the criteria could be hardened, the subjects conceptualised in more appropriate ways, where necessary, and the parts to be played by internal assessment, external assessment and moderation more precisely defined then it was possible that the problems could be overcome. Comparisons were made between external/internal and external/regraded internal assessments. While it was quite impossible to reach precise conclusions backed by classical statistics because of the nature of the grading scheme, the general impression was that there was considerably more agreement than disagreement even although one would expect good agreement anyway when using short scales. So we forged on for a further two years, through 1982 and 1983.

The major drawback with development programmes of this type is that you have to indicate to teachers the requirements for the following year before you have fully analysed and drawn conclusions from the results of the trial in the preceding year. Thus in the second year some steps were made forward, and some back.

Mathematics for example rolled all its assessable elements into one (other than basic skills) and made holistic judgements on a 20 point scale using a Rasch calibrated item bank to fix cut-off points. This worked since it was really resorting to something close to our existing system but it did not satisfy the

philosophy of being able to report in subject terms what a pupil could do. Science continued to confuse itself with a bewildering variety of assessment instruments without getting to grips with a workable conceptualisation of the subject in domain terms. English continued with its four domains of reading, writing, listening and talking but introduced an external listening test to ease the problems of assessment in this area. English also introduced borderline categories in the three point awarding scale to enable a more workable scheme of moderation of assessment to be tested.

Let me now jump forward to this current year of trial, and comment also on other developments which were taking place during the period of the trial which have affected its final design.

As trials were going on a group, of which I was a member, was given the task of investigating Grade Related Criteria as one possible solution to the problems of providing articulated courses and awards, problems which had to be at least partially resolved before the Government would give the go-ahead to the new scheme. There was much scepticism, and it is only fair to say that some still remains, within and without the Group, but nevertheless 9 out of 10 subject groups which were set up to examine the feasibility of Grade Related Criteria within their own areas did in fact report that the idea seemed feasible and recommended it for further study. At the same time the Steering Group postulated the idea of an overlapping awards scheme to parallel the proposed overlapping courses. In such a scheme the award points particularly at the overlapping points would be anchored by Grade Related Criteria as shown:

<u>Original</u>	<u>Proposed</u>	<u>Reported as</u>
Credit	Credit 1	Credit 1
General 1	Credit 2 - General 1	Credit 2
General 2	General 2	General 1
General 3	General 3 - Foundation 1	General 2
Foundation	Foundation 2	Foundation
Course Completed	Course Completed	Course Completed

With some minor modifications this has been accepted by the Government as a key feature of the new scheme.

Following the initial task of investigating Grade Related Criteria the Government provisionally set up Working Parties to devise overlapping courses, and Grade Related Criteria on which awards could be based. Those for English, Mathematics and Science have completed this initial task. In other subjects discussions continue or are just commencing.

A further parallel line of development has involved discussions between the Examination Board and officials of the Government's Scottish Education Department to resolve the philosophy and technical difficulties associated with such a system.

Where do all these developments take us:

First, I must give you the Scottish working definition of Grade Related Criteria. Grade Related Criteria are defined as "broad indicators or descriptions of expected performance in key aspects of a given subject at different levels of award". Their purpose is to provide national guidance to teachers, examiners and the public about the standards applied in the assessable elements of particular subjects. They are in themselves however not self-sufficient but must be presented in summary and extended forms and be read in relation to aims, syllabuses, assessment objectives etc. They are desirable for public communication, course construction, assessment and pupil motivation. They also in the Scottish situation offer a solution to the problems in Dunning differentiation.

Examples of these are given in the reports of the working parties on English, Science and Mathematics. I will however illustrate these by reproducing the Summary and extended Grade Related Criteria for English/Reading and for Science the summary Grade Related Criteria and extended GRC for Handling Information. The Grade Numbers refer to three of the award points on the proposed certificate scale.

(Full sets for English, Mathematics and Science are in the 3 reports which I have brought with me. Regrettably I have only one copy of each but I am happy for anyone to copy the relevant portions or to contact me during the conference so that I may forward a copy on my return to Scotland.)

GRADE RELATED CRITERIA : READING

Applying to all purposes :

SUMMARY

GRADE 6

The candidate will show that he has attempted with some degree of success the main purposes of reading in relation to texts that will for the most part be readily accessible, related to personal interests and experience, dealing with concrete human relationships or containing clearly presented ideas. He will demonstrate a rudimentary understanding and appreciation of the text.

GRADE 5

The candidate will show that he has achieved the main purposes of reading in relation to texts that will be accessible as a whole, mainly related to personal interests and experience, dealing with concrete human relationships or containing clearly presented ideas. He will demonstrate a moderate understanding and appreciation of the text.

GRADE 2

The candidate will show that he has a confident and consistent command of the main purposes of reading in relation to texts that go beyond what is immediately accessible or related to personal interests and experience, some of which will feature unfamiliar, abstract ideas and complexity of structure and tone. He will demonstrate a firm grasp and sensitive appreciation of the text.

EXTENDED

Purpose 1: To gain overall impression, gist of a text.

GRADE 6

The candidate will accurately identify * the main concern(s) of the text.

GRADE 5

The candidate will accurately identify or state ** in his own words the main concern(s) of the text.

GRADE 2

The candidate will make a clear and concise statement of the main concerns of the text and, where appropriate will show a clear perception of their inter-relationships.

* Identify implies underlining, ringing, numbering, ticking, or otherwise 'marking' the text - as well as writing or talking.

** State implies writing or talking.

Purpose 2: To obtain particular information from a text.

GRADE 6

The candidate will accurately identify individual items of information retrieved from areas of the text that have been clearly defined.

GRADE 5

The candidate will state accurately, not necessarily in his own words, individual items of information retrieved from the text.

GRADE 2

The candidate will state accurately in his own words and, where appropriate, collate items of information retrieved from the text.

Purpose 3: To grasp ideas or feelings implied in a text.

GRADE 6

The candidate will draw an acceptable inference from a key statement in the text, and/or identify a statement that would permit a given inference to be drawn.

GRADE 5

The candidate will draw an apt inference from a key statement in the text, and/or identify a statement that would permit a given inference to be drawn.

GRADE 2

The candidate will draw an apt inference from a key statement or statements in the text, and will be able to substantiate this by reference to appropriate textual evidence.

Purpose 4: To evaluate the writer's attitudes, assumptions and arguments.

GRADE 6

The candidate will make a simple, intelligible comment on some clearly defined aspect of the writer's point of view, and will attempt to justify it primarily by reference to his own experience.

GRADE 5

The candidate will make a simple, intelligible comment on some clearly defined aspect of the writer's point of view, and will attempt to justify it by reference to his own experience and knowledge.

GRADE 2

The candidate will make a considered comment on some aspect of the writer's point of view, and will show some skill in justifying it by a degree of argument that makes reference to his own experience and knowledge and, where appropriate, textual evidence.

Handling Information

Set B

Summary GRC

Grade 2

To be awarded this Grade a pupil has demonstrated consistently the ability to extract and collate information from a variety of self sought sources and present information in appropriate forms with a high degree of neatness, skill and accuracy. The pupil has shown the ability to draw valid conclusions in new situations, express these conclusions with clarity and go on to form a hypothesis based on the conclusions.

Grade 5

To be awarded this Grade a pupil has demonstrated the ability to extract and collate information from a number of suggested sources and present information in appropriate forms with some degree of neatness, skill and accuracy. The pupil has shown the ability to draw valid conclusions in related situations and to suggest what might happen in such situations.

Grade 6

To be awarded this Grade a pupil has demonstrated the ability to extract and collate information from simple given sources and present information in appropriate forms normally by completion and with guidance. The pupil has shown the ability to draw conclusions in simple situations, and to suggest what might happen in a similar situation.

HANDLING INFORMATION - EXTENDED GRC

Grade 6

A pupil awarded this Grade will have demonstrated the ability to:-

locate information by using the index of a book;

obtain relevant information from a given source, to include a passage of writing, picture, key and flow chart;

extract relevant data from a given source to include a table, diagram, key, flow chart, bar graph, pie chart and line graph;

present information in oral and written forms;

complete tables of data when the headings are given;

draw and label diagrams;

complete flow charts, keys, bar graphs and line graphs;

draw an appropriate conclusion from simple scientific information;

use a conclusion to suggest what might happen in a similar situation

Grade 5

In addition a pupil obtaining this Grade will have demonstrated the ability to:-

obtain and collate information from a number of suggested sources;

construct tables of data when no headings are given;

draw and label simple pie charts;

construct bar graphs and line graphs given the scales and the axes;

draw appropriate conclusions from scientific information;

use conclusions to suggest what might happen in related situations.

Grade 2

In addition a pupil obtaining this Grade will have demonstrated the ability to:-

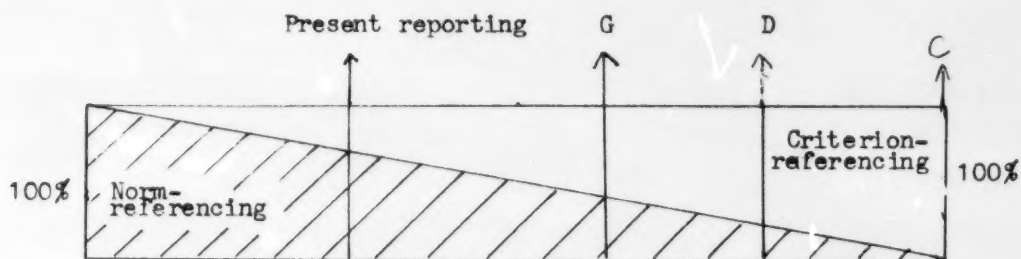
distinguish the main theme of a piece of scientific writing;

extract and collate information from a variety of self sought sources.

construct pie charts and line graphs when scales and axes are not given;

use conclusions to form a hypothesis or a generalisation.

It might also be useful at this point to consider where Grade Related Criteria fit into the broad scheme of things in assessment terms. There is a notion that norm-referenced certification is at the opposite end of a spectrum from criterion-referenced certification, though the view is also held that there is an element of each in the other. I tend to agree with the latter when we are considering certification of school pupils over a broad educational curriculum, though it is obvious that the emphasis can be pushed one way or the other. However, if we consider the spectrum model, as illustrated below we can see that criterion-referencing in the situation we are considering tends to become domain-referencing through practical considerations. It is impracticable to handle and assess every skill process or criterion on its own. This I have indicated at an arbitrary point D. Summary GRC attempts to move from norm-referencing to criterion-referencing but in terms of reporting achievement. This is indicated at point G. But as I have already stated, summary GRC are useless on their own. They require further subdivision into extended form before they are of practical use for assessment purposes. And what we find from a purely empirical approach such as this is that the extended GRC are very similar, if indeed not the same as the domain-referenced approach. Thus Grade Related Criteria in reality becomes a method of reporting the results of an assessment approach which is to all intents and purposes domain referenced. This is particularly significant in a national certification process since reporting of criterion-referenced assessment as a direct output has been considered by many to be too unwieldy and too complex to apply to national certification on a large scale. However, profile reporting in this scheme would produce output at point D.



To return now to English, Mathematics and Science. The final position with these subjects regarding domains is as follows:

- | | |
|----------|---|
| English: | Reading, Writing, Talking and Listening |
| Science: | Knowledge and Understanding, Handling Information, Practical Skills and Problem Solving |

Mathematics: Interpreting/Communicating Information, Selecting a Strategy,
Processing Data, Handling Practical Situations, Problem
Solving

I will now give a brief outline of the proposed system as a result of the pilot work. Space and time mean that I can only give you a flavour of the system but I hope it will suffice for any of these present who wish to pursue discussion of the proposals during the conference or at a future date.

The strategy is as follows. Grade Related Criteria have been defined for at least 3 of the 7 points in the awarding system which covers the whole ability range. Point 7 "course completed" represents failure to reach minimum competence and so is self-defined. It is intended that criteria will also be defined for the remaining points on the scale in the near future.

These Grade Related Criteria are specific to each domain and to achieve a grade in that domain the pupil must achieve the criteria listed in the extended Grade Related Criteria. In cases where internal assessment is permitted (and we hope it will be eventually permitted at all levels,) teachers will use the criteria to grade the candidate. In other cases teachers will be asked anyway to estimate a grade for the candidate in each domain. Wherever feasible an external assessment will be made of the candidate in each domain. In science for example this will be in all domains but practical. There will be three papers one on Knowledge and Understanding, one on Handling Information and one on Problem Solving. The Practical Skills based on some 50 specific criteria will be internally assessed. In English there will be the following arrangement:

Reading:	Folio + external test
Writing:	Folic + external test
Listening:	External test only
Talking:	Internal assessment only

External papers will be either directly graded during marking, or cut-off points will be determined in relation to the criteria if the nature of the test makes this a simpler arrangement.

Where there are external and internal measures of the same domain to the same criteria any small discrepancies will be reconciled, normally by giving the benefit of the doubt to the candidate. Where there is a large discrepancy school evidence will be called for and the examiners will use this to resolve the discrepancy one way or the other. Where there is one measure only the grade will stand. Inspection procedures on a whole school basis may be introduced to ensure that there is no abuse of the system.

It is our firm hope that the domain performance will be reported on the certificate in the form of an attainment profile and a new certificate has been designed to that end.

Aggregate awards will also be reported but these may allow trade-offs between varying performances in different domains - this has yet to be resolved. Our first certificates of this nature should be produced in Scotland in 1986.

Examples of this year's trial papers (10 copies of each) are on display. There is a tape for the listening test in English of which I have one copy here. I should caution you that they do not all conform to the pattern described since these were designed before the final pattern was decided. Science, for example has only two papers, though these cover the three domains mentioned before. Papers in the operational situation will however conform to the framework mentioned.

I hope you have found sufficient in this development to interest you at least in some of its aspects, whether it be the pitfalls or the successes.